

利己性と利他性と社会的ディレンマ

高 田 洋 一 郎

1 利己性の問題

人間を含む生物の利己性の問題は古くして新しい。Hobbes は 1651 年、その著『リヴァイアサン』(*Leviathan*) のなかで、人間は本来、利己的な存在であり、利己心に基づく競争は必然的に「各人の各人に対する戦争」(*warre of every man against every man*) を生むといい、それを防ぐための権力機構として専制国家の存在を基礎づけようとした。それから百数十年たって Adam Smith は『諸国民の富』(*An inquiry into the nature and causes of the wealth of nations*, 1776) のなかで、個人の利益を求める利己的な行動は「見えない手」(*an invisible hand*) に導かれて本人の意図しなかった社会の利益・公共の幸福 (*interest of the society, public good*) を促進することがしばしばあると主張した。Hobbes と Adam Smith の主張は一見、正反対のようであるが、利己心に基づく競争が人間の本性であるという認識において両者は共通している。ただ、Hobbes がそこから生じるのは果てしない闘争というディレンマ（個人の利益を求める行動が、かえって個人の利益を損うという意味でのディレンマ）であって、このディレンマを解消するには強力な外部権力の介入を必要とすると考えたのに対して、Adam Smith はむしろそこから社会全体の調和と利益が結果すると考えた点が違っている。

どちらの認識が正しいのであろうか。その後の経済学と政治学の発展は、何れの認識もある意味では正しく、しかし厳密には正しくないことを示しているようである。19 世紀後半、Edgeworth と Walras にはじまり、約 1 世紀を経て今世紀半ばすぎ Arrow らによってほぼ完成をみた一般均衡理論

は、精巧な数学的モデルを使って、Adam Smith の「見えない手」の働きを証明することに成功した（一般均衡理論の解説書は内外に数多いが、最近の成果の体系的叙述としては Arrow & Hahn 1971, また Edgeworth と Walras の発想の現代的解釈としては「Edgeworth と Walras の主題による変奏曲」という魅力的な副題のついた Hildenbrand & Kirman 1976 がある）。すなわち、利己心の自由な発露と競争を保証する市場が存在すれば、各経済主体が利己心の最大の満足を求めて行動するとき、市場の価格機構が〈見える〉導きの手として働いて、その結果、自動的に社会全体の資源の効率的配分が実現する、というのである。ここで効率的というのはいわゆる〈パレート最適〉 (Pereto optimal) 状態であって、誰かの利益を損うことなしには誰の利益も増すことができないような状態をいう。パレート最適でない状態は社会全体からみて改善の余地がある状態であるから、パレート最適性は公共の幸福のための必要条件であるといえよう。すなわち一般均衡理論は、利己心に基づく個人の自由な競争が公共の幸福に必要な条件を自動的に実現することを証明したという意味で、Adam Smith の主張を裏付けたといえる。

しかし同時にこの理論は、Adam Smith の主張の限界をも明らかにした。市場はしばしば自由ではないが（独占や寡占）、それを自由にしてパレート最適性を実現するには外からの権力の介入（たとえば独占禁止法の制定とその強制）を必要とする。さらに深刻な問題は、市場の存在し得ない財やサービスに対しては価格機構が機能しないことである。いわゆる〈外部性〉 (externalities) と〈公共財〉 (public goods) の問題であるが、これらに対してパレート最適性を保証するにはますます多くの公権力の介入を必要とする。さらにまた、パレート最適性は公共の幸福の必要条件ではあっても十分条件ではないという問題がある。極端な場合、一国の富が一人に集中した状態は、いかなる富の再配分もその一人の利益を損うから定義によってパレート最適であるが、これを公共の幸福と考える人は少ないだろう。つまり公共の幸福というとき、パレート最適性という効率性の基準のほかに、何らかの

意味で〈公正〉を測る基準が必要だということになるが、一般均衡理論の体系のなかにそのような基準は存在しない。むしろ Arrow の古典的論文 “*Social choice and individual values*” (1951) が鋭く示したように、個人の満足を合成して公共の幸福を導く手続そのものに大きな論理的困難があるのであるが、これはいわゆる〈投票のパラドックス〉として、民主政治の基盤にもかかわる重大な問題となっている。

以上のように、経済学の現在までの展開は Adam Smith の世界を次第に Hobbes の世界に近づけていったといえよう。同時にそれは経済学と政治学の接近であり、この接近のなかで、Adam Smith 的調和と Hobbes 的ディレンマとを同じ楯の両面と見る可能性が暗示されているように思われる。

ところで、以上のような利己心に基づく競争、また競争が生む調和とディレンマは、人間のみに特有な現象であろうか。Adam Smith の『諸国民の富』からさらに 80 年あまりを経て出版された Darwin の『種の起源』(*The origin of species by means of natural selection*, 1859) は、これが人間のみならず生物一般の原則であることを強く示したといえよう。ただし Darwin は生物の利己性が競争を生むというより、むしろ生物界に普遍的に存在する〈生存競争〉(struggle for existence) が生物の利己性(世代を通じて自己と自己の子孫の保存を求める遺伝的傾向)を必然ならしめると考えたのである。生存競争とは世代を通じての生物間の生き残り競争であり、その勝敗は生物個体の生存率と繁殖率によってきまる。生存率が高いほど繁殖機会も増すから、両者をあわせて、生物個体が次世代に残す子孫の数の多少が生存競争の勝敗を分けるといってもよい。この、生物が次世代に残す子孫の数(通常、親を数えたのと同じ齢で子の数の平均を数えるが、後の例が示すように孫の数を数えた方が適当な場合もある)を〈ダーウィン適応度〉(Darwinian fitness) と今日いうが、生存競争とは要するにダーウィン適応度をめぐる競争である。一般に、産まれてくる子や卵の数は親の数より遙かに多いのに生物の総数は比較的安定している——この事実から Darwin

は生存競争の普遍性を推定した。いま一つ Darwin の進化論の基礎になったのは遺伝する変異の存在である。この、生存競争の普遍性と遺伝する変異の存在という二つから、避けることのできない結論として導かれたのが自然淘汰による進化の理論であった。生物はそれぞれの環境でダーウィーン適応度をめぐって競争する。ダーウィーン適応度を高めるような遺伝的変異を持つ個体はより多くの子孫を残し、子孫は親と同じ遺伝的変異をもつことが多いから（ただし後述のように、この関係は一般に簡単でない）、世代を重ねるにつれて、適応上有利な遺伝的変異が繁殖集団のなかに蓄積してゆくことになる。このようにして生物集団は次第に、それぞれの環境に適した方向にその遺伝的構成を変えてゆき、その結果、種の分化が生じる。以上が Darwin の考えのあらましであるが、ダーウィーン適応度の上昇とは生物が世代を通じて計る自己保存の増大であるから、これを簡単に生物の利己性への傾向と呼ぶとすれば、進化とは生物が遺伝を通じて利己性を獲得する過程である、ということになる。この意味で〈最適者生存〉とは〈最大利己者生存〉に他ならない。以上は生存競争と遺伝する変異を認める限り避けることのできない結論、あるいは自明の理であるといっていよい。

あまりにも自明であるために、かえってトートロジーだという批判が現在でもあとを絶たない（たとえば Macbeth 1971）。実際、〈最適者生存〉を「最も良く生き残る者が生き残る」と解釈すれば、これは同語反復以外の何者でもない。しかし、ここでいう生き残りにには、個体の生存だけでなく繁殖を通じての子孫の生き残りが含まれること、さらに〈良く生き残る〉という性質がどのように子孫に伝えられるかについての遺伝法則が関係していることを考えれば、最適者生存が決して同語反復でないことは明らかだろう。事実、適者が多くの子孫を残すことによって、適者であるという性質がどのように集団に蓄積してゆくか、それを正確に論ずるには正しい遺伝法則を必要とするが、それは Darwin にはまだ与えられていなかった。Mendel による遺伝法則の発見は 1865 年であるが、1900 年の再発見まで世の知るところとならなかったことは周知の通りである。したがって Darwin の適者生存

の論理に、その分、曖昧さが残っていることは止むを得ない。メンデル法則再発見後 1930 年代までに Fisher, Haldane, Wright の 3 人によってつくられた集団遺伝学は、数学的モデルによって遺伝法則と進化論の関係を定量化することに成功した (Fisher 1930, Wright 1969, 1977, 1978. Fisher の古典的著作は、あとで若干触れるように、今日読んでも啓発されることが多い。Wright の大著は最近までの理論と実証の集大成である)。今日、集団遺伝学などで用いられている「進化とは繁殖集団における (対立) 遺伝子頻度の変化である」という定義は、Darwin の論理を数学的に明確化する上で大いに役立ったし、今後も有効に機能することが期待される (ただし大進化を論ずるにもこれで十分かという疑問が残らないではない)。

集団遺伝学の成果を取り入れて体系化された進化論は〈進化総合説〉(ネオ・ダーウィニズム)として一時、正統性を確立したかにみえたが、現在、ミクロとマクロの両面から、〈中立説〉(Kimura 1968)と〈断続平衡説〉(Eldredge and Gould 1972)という批判が登場している。前者は最近のめざましい分子進化学の発展 (たとえば木村 1984 参照) の中から生まれたもので、分子レベルでは淘汰上有利でも不利でもない (中立的) 変異 (遺伝子コードの指定する蛋白質のアミノ酸配列の変異、さらに一層直接的に DNA 塩基配列の変異) が数多く存在することを指摘することによって、自然淘汰万能の総合説への批判を提起しており、後者は主として化石生物学上の証拠から、進化が漸進的には進まないことを主張して総合説と対立している。このような最近の進化生物学の展開は、広く生物科学の将来を考える上でも重要であり、筆者のような生物学専攻でない者の関心をも強く惹く話題であるが、中立説にせよ断続平衡説にせよ、自然淘汰を主因とする進化という Darwin の基本構想を否定するものでないことは、これらの説の主唱者自身が説く通りである。むしろ Gould などは、Darwin の進化論が今日なお正しく理解されていない (たとえば進化と進歩の混同は今日でも稀でない) ことを嘆いて、Darwin の発想に立ち返って考えることの重要性を指摘している (Gould 1977)。

以上のように、Darwin 以来現在にいたるまでの進化論の考えでは、生物の利己性はほとんど自明の事実であるといっていよい。長い進化の過程で生物はさまざまな利己的遺伝形質、すなわちそれぞれの環境でダーウィン適応度を高めるような解剖学的、生理学的、行動的特性を蓄積してきたはずであり、非利己的な形質はたとえ一時存在したとしても自然淘汰によって除かれているはずである。こうした考えを一層明確に表わしたのが〈利己的な遺伝子〉(selfish gene) という言葉であろう。上記の「対立遺伝子頻度の変化としての進化」、という捉え方のなかに、〈利己的〉(自己保存的)なのは究極的には遺伝子であるという考え方がすでに含まれている。進化が遺伝子頻度の変化であるなら、進化の結果は自己保存に最も巧みな対立遺伝子の勝利として現われるだろう。遺伝子は、発生の過程でその暗号コードによって生物の身体を作り、繁殖を通じて次の世代に自らのコピーを伝えるから、身体の手作り方の如何で自らの命運が左右される。そこで当然、遺伝子をして自らの保存に有利な身体を作らしめるよう淘汰圧が働くはずだということになる。ここで身体というのは単に生物の形態的特徴ばかりでなく、神経系や内分泌系も含むから、行動や心の働きも含む広い意味での身体が遺伝子にとって自己保存の手段であるといっていよい。別の言葉でいうと、生存競争の本質は繁殖集団における対立遺伝子間の生き残り競争であるということになる。〈利己的な遺伝子〉というのは、Dawkins がその広く読まれた行動生態学の啓蒙書 (*The selfish gene*, 1976) につけた標題であるが、このような考え方が特に効力を発揮するのは、遺伝子の自己保存性が個体の自己保存と反対の方向に働く場合である。遺伝子が本当に〈利己的〉であれば、生物個体の自己犠牲を通じて遺伝子のよりよい自己保存が達成される場合には、自己犠牲という、一見非適応的な行動が進化しうることになるからである。Hamilton の〈血縁淘汰〉の理論はそうした可能性を示す一例であるが、これについては後に述べる。

Dawkins は上掲書で生命の起源にふれて、DNA 様の自己複製子 (replicator) を考えている。自己複製子に複製の過程でランダムなコピー・ミス

（〈突然変異〉）が生じ、その結果の〈変異〉に自然淘汰が働けば、自己複製子の〈進化〉が起きることになる。自己複製子が進化の過程で創造した〈乗物〉が生物の身体であり、身体の複雑化の結果として神経系が生まれ、遂に人間の心の機能までが生じたとみることができる。〈進化する〉自己複製子は定義によって〈利己的〉である。つまり〈利己的な自己複製子〉とは〈コピー・ミスと淘汰によって進化する自己複製子〉と同義であるといっていよい。

このように考えると、生物の利己性とは生命の起源にまでさかのぼる生物進化の原則であるといわざるをえない。また Hobbes や Adam Smith の指摘した人間の利己性の問題も、人間だけの問題としてではなく、生物全体の問題として考えた方がよいことになる。と言っても、人間の利己心がすべて進化の産物であるというわけではない。第一に上述のように、遺伝子の利己性が直ちに行動の利己性を生むわけではないという事情がある。第二に、しばしば指摘されることであるが、人間の行動は多分に学習と文化の産物であって、遺伝と進化の法則だけの支配下にあるのではない。もちろん動物にも学習能力はあるし文化の萌芽もみられるが、人間ほど徹底した文化的環境を創造した生物は他にいない。文化という人為的環境のなかで生活する文明社会の人間は、直接自然から淘汰の圧力を受けることもなく、ダーウィン適応度をめぐって競争することもないようにみえる。したがって人間の行動を理解するには、生物学的進化とともに文化的進化を考える必要があることは言うまでもない。しかし、ここで注意を要するのは、少なくとも人間文化の発生期には生物進化の法則が強く働いたであろうこと、また文化の維持と発展もおそらく生物進化の法則と無関係ではなかったであろうということである。このように考えるとき、Hobbes や Adam Smith の指摘した利己心に基づく競争と、それがもたらす調和とディレンマの問題を始めから人間固有の問題と限定してかかるのではなく、広く生物一般の問題として考えるのが生産的ではないかと思われるのである。そこでまず、Darwin 自身がこの問題をどのように考えていたかを見ることとしよう。

2 Darwin と利他行動

上述のように、Darwin の進化論にとって生物の利己性はほとんど自明の理であったが、生物界には一見、利己性に反するような行動も存在する。たとえば社会性昆虫におけるワーカーは不妊の雌であるが、不妊であることによってワーカーの個体としてのダーウィン適応度は低下している。また生物界には競争と同時に協力行動もあまた存在するが、そのなかには一見、個体のダーウィン適応度を低下させているような行動（たとえば鳥類や哺乳類の警告声）もある。とくに人間の社会は、個人の利己心の抑制による協力行動に支えられて成立しているようにみえることが多い。こうした利他行動（個体のダーウィン適応度を低下させる行動）の存在は Darwin にとっても大きな問題であった。『種の起源』のなかで Darwin が社会性昆虫の進化の説明に用いたのは、現在の用語でいえば〈群淘汰〉(group selection) の理論と血縁淘汰 (kin selection) の理論であるが、「人類の起源と性淘汰」(*The descent of man and selection in relation to sex*, 1871) では人類の道徳性の起源について鋭く問題の本質を突いた議論を展開している。それは今日見ても驚くほどの洞察である。以下、上掲書の第5章「原始時代と文明時代における知的能力と道徳的能力の発達について」から、関連する部分を引用する。

- 1) When two tribes of primeval man, living in the same country, came into competition, if (other circumstances being equal) the one tribe included a great number of courageous, sympathetic and faithful members, who were always ready to warn each other of danger, to aid and defend each other, this tribe would succeed better and conquer the other.

ここで Darwin が取り上げている人類の道徳的能力は勇気、思いやり、誠実さなどであるが、これらの特性を直ちに利他行動とみなすことには問題があるかもしれない。たとえば互いに誠実であった方が相互の利益になる場合

もあるからである。しかしあとでも述べるように、相手が誠実であるとき自分は不誠実であることによって一層利益を受ける場合があるが、そうした場合でも誠実であることは自己の利益を犠牲にして相手の利益を計るという意味で利他行動であるといつてよい。このような利他行動がどのようにして進化したのであろうか。Darwin が上の引用で用いた説明は前記の〈群淘汰〉、すなわち同種生物の分集団の間で働く自然淘汰であった。利他行動は個体の適応度を低下させるから、通常は自然淘汰によって排除され、集団の支配的な特性になることはない。しかし、利他者を含む集団と含まぬ集団があって、両者が生殖的に隔離され、しかもその間に生存競争が存在したとすると、利他者を含む集団の方が含まない集団より適応上有利であるから、前者が競争に勝ち後者は淘汰され、やがては利他者が集団全体を支配するようになる。これが群淘汰による利他行動の進化である。ただしこのようなことが実際起きるためには、集団内で利他者が淘汰されてしまう前に、集団間の淘汰が完了しなければならない。つまり集団内で個体に働く群内淘汰にくらべて、集団間で働く群間淘汰の淘汰圧が十分強いことが必要であるが、この点に群淘汰説の難点がある。一般に群内淘汰にくらべると群間淘汰はずっと力が弱いと考えられるからである。こうした理由から、現在の進化生物学でも群淘汰説を認める者と認めない者とが分かれている。概して言えばイギリスの行動生態学者には認めない者が多いが、アメリカの社会生物学者や集団遺伝学者には認める者もいるようである。何れにせよ、群淘汰の成立する条件が相当きびしいことは確かで、この問題の解明は今後に残されているが、注目されるのは Darwin がすでに群淘汰説の問題点に気づいていたことである。それは上記の引用の十数行あとに続く次の文章から察せられる。

- 2) But it may be asked, how within the limits of the same tribe did a large number of members first become endowed with these social and moral qualities, and how was the standard of excellence raised?... (中略) ...He who was ready to sacrifice his life, as many a savage has been, rather than betray his comrades, would often

leave no offspring to inherit his noble nature.

引用文の前半は、現代風にいえば、厳しい個体淘汰の下で、そもそもどのようにして利他性が集団内で支配を獲得しうるかということであり、後半の「仲間を裏切るより…むしろ進んで自己の命を犠牲にする者は、しばしばその高貴な性質を遺伝する子孫を残さないであろう」という文は、まさに問題の本質（利他行動自体によって利他行動が淘汰されるということ）を的確に言い当てている。

それではこの矛盾はどう解いたらよいのであろうか。Darwin が上の引用文に続く文章で「明確に分析するには余りにも複雑な状況であるが」と断りながらも、一つの可能性として提案するのは次のようなメカニズムである。

- 3) Although the circumstances, leading to an increase in the number of those thus endowed within the same tribe, are too complex to be clearly followed out, we can trace some of the probable steps. In the first place, as the reasoning powers and foresight of the members became improved, each man would soon learn that if he aided his fellow-men, he would commonly receive aid in return. From this low motive he might acquire the habit of aiding his fellows; and the habit of performing benevolent actions certainly strengthens the feeling of sympathy which gives the first impulse to benevolent actions. Habits, moreover, followed during many generations probably tend to be inherited.

明らかにここで述べられているのはお返し (aid in return) 期待の利他行動であり、本来、利己性と矛盾しない、いわばみかけの利他性である。つまり、もともと利己的な動機に基づくという意味で Darwin も low motive と言ったのであろう。注目されるのは、これが 100 年後の 1971 年、イギリスの Trivers によって非血縁個体間の利他行動を説明するために提案された〈互惠的利他行動〉(reciprocal altruism) の理論と同じ発想を示している

ということである。この発想は人間社会についてはむしろ常識的であろう。実際、一見すると無私の利他行動が実は暗黙のお返し期待に基づいて成立していることがしばしばある。本人がお返しなど期待しないと思っているときでも、相応のお返しのないとき立腹することがあるのは意識下に期待があったからであろう。お返しの意識を抑圧することが適応的である可能性すら存在することを考えると、互惠的利他行動は見かけ以上に人間社会に普遍的に存在するのかもしれない。また人間社会以外にも、たとえばアヌビスヒヒ *Papio anubis* の雄の協力関係について Packer (1977) は、互惠的利他行動であることを示す観察を報告している。

互惠的利他行動論の一番の難点は、すでに述べたように、協力の相互性を維持するメカニズムの問題である。相手から協力を受けることによって利益を得た個体は、相手に協力を返さないことによって一層大きな利益を受けることが多い。そうだとすると利己性に基づく協力行動は一時的には成立しても安定的に維持されないことになる。つまりお返し不履行 (cheating) あるいはただ乗り (free riding) を防止するメカニズムの進化の説明なしでは上記の理論は利他行動の説明として十分ではないのである。Darwin はこの点に触れていないが、100年後の Trivers の理論もこの点について十分説得力のある説明をしていない。実はこのような、「相互協力によって利益を受けるが、相手の協力に便乗することで一層大きな利益を得るために、相互協力が安定的に成立しない」という状況の普遍性が近年になって各方面から指摘され、さらにこのような状況で非協力が一般化すると結局、相互協力にくらべても損な結果になるというディレンマがしばしば問題にされるようになったが、この点については後述する。

Trivers の互惠的利他行動論は血縁関係のない個体間の協力行動を説明するための理論であるが、血縁個体間の協力行動については Hamilton (1964) の〈血縁淘汰〉(kin selection) の理論がある。驚くべきことに Darwin はこの理論の可能性まで示唆しているのである。前記 1) の引用に先立つ部分で Darwin は人類の知的能力の進化について論じているが、知的能力の進

化には道德性の進化のような理論的困難はないと彼は考えた。何故なら、知的能力は一般に生活技術の改善を通じて個体の適応度を上昇させる——つまり知的能力の高い個体ほど良く生きのび、多くの子孫を養育できる——からである。そうした論を述べたあとに次の文が続く。

- 4) If such men left children to inherit their mental superiority, the chance of the birth of still more ingenious members would be somewhat better, and in a very small tribe decidedly better. Even if they left no children, the tribe would still include their blood-relations; and it has been ascertained by agriculturists that by preserving and breeding from the family of an animal, which when slaughtered was found to be valuable, the desired character has been obtained.

つまり知的素質のすぐれた者が子どもを残せば、子孫の中に知的素質の一層すぐれた者が増えてゆくが、たとえ自らは子どもを残さなくとも、そのすぐれた素質は血縁者（きょうだい、おい、めい、いとこなど）を通じて後の世代に伝わってゆくであろうというのであるが、このことを厳密に定量化して理論化したのが Hamilton の理論に他ならない。

以上のように、利他行動の進化をめぐる現在の議論はすべて Darwin にその萌芽があると言って過言でないのであるが、以下その現代的展開としてまず Hamilton の理論を取り上げる。

3 Hamilton の血縁淘汰の理論

Hamilton の理論は簡単にいえば、遺伝子は直接の子以外にも血縁を通じて後の世代に伝わってゆくということに注目した理論である。通常の有性繁殖を行う 2 倍体生物では、親のもつ相同遺伝子は減数分裂の際に別れて、それぞれ独立に配偶子に入ってゆくの、親の遺伝子の半分が子に伝えられることになる。同じことを、親と子が共通祖先に由来する遺伝子を共有する確

率（同祖遺伝子共有確率）は $1/2$ であるといってもよい。ただしこの場合、親と子の共通祖先とは親自身のことであるが、こうした回りくどい言い方をしたのは親子以外の血縁関係にも同じような確率を考えたいからである。たとえば完全同胞（両親を共通にもつぎょうだい）について考えてみよう。いまぎょうだいの一方が持つある遺伝子に着目する。ぎょうだいの至近の共通祖先は両親であるから、その同じ遺伝子（のコピー）は両親を通じてぎょうだいの他方に伝えられている可能性があるが、その確率はどの位だろうか。問題の遺伝子が父親に由来する確率は $1/2$ であり、父親に由来するときそれがぎょうだいの他方にも伝えられている確率は上記のように $1/2$ だから、結局、父親経由でぎょうだいに伝わっている確率は $1/2 \times 1/2 = 1/4$ である。母親経由の確率も同様に $1/4$ である。したがって両親の何れかを經由して同一遺伝子（のコピー）をぎょうだいが共有する確率は $1/4 + 1/4 = 1/2$ となる。もし両親が血縁であれば（たとえばいとこ婚の場合）、さらに両親の共通祖先にさかのぼってその分を加えなければならないが、血縁でない両親の子の間での同祖遺伝子共有確率は $1/2$ であり、親子間の値と同じ値になる。以下、この確率を〈血縁度〉(relatedness) と呼ぶ。半同胞の血縁度は上述から明らかなように $1/4$ であり、いとこ間では $1/8$ であることが容易に計算できる。

ここで血縁度を定義するのが〈同祖〉遺伝子共有確率であって、単に遺伝子共有確率ではないことに注意を要する。この点の誤解が多いようであるが (Dawkins 1979), 単に 2 個体が共有する遺伝子の割合であれば同種生物間では極めて高い値になるし、異種間でも相当な値になることがある。何故、同祖と限定する必要があるのかを厳密に論ずることは省略するが、次のように考えれば直観的には理解できよう。一般に進化において問題になるのは、突然変異によって生じた対立遺伝子がいかにして繁殖集団の中に拡まってゆくかということであるが、突然変異の発生率は一般に低いので（1 遺伝子座 1 世代あたり $10^{-4} \sim 10^{-6}$ のオーダー）、同じ突然変異が同時多発的に集団中に発生するとは考えにくい。したがって集団の中のある個体のある遺伝子座

にたまたま突然変異によって生じた対立遺伝子が、どのようにして世代を超えて伝播してゆくかを考えればよいことになる。いま問題なのは利他行動の遺伝子（といっても単一の遺伝子である必要はない。遺伝子複合であっても基本的には同じであるが、ここでは簡単のため単一遺伝子として考える）であるが、利他行動遺伝子がある個体 X と別の個体 Y との間で共有されるのは、 $X(Y)$ の突然変異が $Y(X)$ に伝えられたか、あるいは X と Y の共通祖先の突然変異が X と Y の両方に伝えられたかの何れかによる。そこで自分自身も自分の祖先に加えて考えれば、両方の場合をあわせて〈同祖遺伝子〉の共有と言ってよいことになる。

さてここで、利他行動を次のように定義する。「利他行動とは、個体が自分の適応度を減らして他の個体の適応度を増す行動である。ただし適応度とはダーウィン適応度をいう」。この定義は客観的であるが、どの程度実用的かという問題がある。この定義によってある行動が真に利他行動か否かを判断するにはダーウィン適応度の変化を調べなければならないが、ある行動によって個体の生存率や妊（稔）性がどれだけ変化したかを正確に推定することは一般にかなり難しい。しかし現に適応度を測る試みがいくつか存在するし、測定法の改善も今後に期待できるので、ここでこの問題に立ち入ることはしない。残る疑問は上の定義が人間にも適用できるかということであろう。人間も生物とはいえ、人間の適応度を生存率や妊性のみで測れるかという疑問が当然生じるであろうが、この問題を考えてゆくとどうしても人間の文化の問題に触れることになるので後廻しにして、ここではただ人間も進化の産物である以上、ダーウィン適応度は人間行動の理解にとっても必要な（十分ではないが）条件であることを指摘するに止めておこう。

さて、ある行動による個体の適応度の減少（その行動をとらない場合と比較しての）を c であらわし（行動の cost の意）、行動の向けられた相手の適応度の増加を b であらわすと（ b は benefit の意）、利他行動とは $b > 0$, $c > 0$ なる行動であると定義できる。利他行動によって自分の子どもは c だけ減るが、その代り相手の子どもは b だけ増える。いま自分と相手の間の血

縁度（同祖遺伝子共有確率）を r とすると、利他行動遺伝子を持つ相手による子どもの増分（の期待値）は rb である。そこで遺伝子の立場で考えると、利他行動の仕手と受け手で利他行動遺伝子を有する者のつくる子どもの数は、利他行動によって $rb - c$ だけ増す（ $rb - c$ がマイナスの場合は減る）ことになる。子どもの数が増えれば一般に子どもの伝える遺伝子の数も増えるだろう。そこで結論として「 $rb - c > 0$ のとき、利他行動によって利他行動遺伝子の頻度（相対頻度）が増す」という規則が成り立つ。進化とは繁殖集団における対立遺伝子間の頻度の変化であるから、これで利他行動の進化が説明できたことになる。また前述の Darwin の疑問——「進んで自分の命を犠牲にする者が、いかにしてその高貴な性質を遺伝する子孫を残すか」という疑問に対する一つの答が得られたことになるが、それが可能になったのは〈利己的な遺伝子〉の立場に立ったからである。

上の規則を血縁淘汰に関する〈Hamilton の規則〉と呼んでいる。Hamilton 自身は上の不等式よりも複雑な〈包括適応度〉 (inclusive fitness) という概念を使って説明しているが、Grafen (1982, 1984) が指摘したように、包括適応度という概念は誤解されやすく、また計算上の困難があって、むしろ上の不等式の方が使いやすい。なお血縁淘汰というのは Maynard Smith の命名であるが、直観的に分かりやすいので広く用いられている。

そもそも真の利他行動であれば、利他行動自体によって利他行動遺伝子の頻度が減るから利他行動は進化しえないはずだ、というのが Darwin の提起した疑問であったが、Hamilton の理論によって、利他行動が血縁間で生じる場合には、血縁度 r 、利他行動が相手に与える利益（適応度増加） b 、自分が蒙る損失 c （適応度減少）の三者の間の関係によって、利他行動の進化する場合のあることが示されたことになる。

それでは生物の利他行動は、実際に Hamilton の規則に従って進化したのであろうか。この点の実証は目下進行中であるが、 r と b と c の実測上の困難もあって、確定的な答を得るにはなおしばらく時間がかかりそうである。ここでは詳しい論評は省略し、興味ある例を一、二あげるに止めたい。

Darwin 以来、利他行動として最も注目をひいてきたのは社会性昆虫におけるカースト（女王、ワーカー、兵隊などの役割分化）の存在であった。たとえばワーカーは通常不妊の雌で自分は子をつくらず、女王の産む弟妹の養育などで一生を終わる。自分の子をつくらないのであるから、定義によってワーカーの行動のほとんどは利他行動であるが、このようなカーストがどのようにして進化してきたのであろうか。Darwin の先駆的な示唆から約 100 年を経て、Hamilton (1964, 1967) の理論によって血縁淘汰による説明の可能性が生じた。しかしそれを述べる前に、社会性昆虫の多くがもつ特徴的な繁殖様式に触れる必要がある。社会性昆虫のほとんどは膜翅目に属する。例外として等翅目のシロアリがあり、また最近日本で、半翅目のアブラムシに兵隊カーストの存在することが発見されたが (Aoki 1977, 青木1984), ここでは膜翅目の社会性昆虫（ハナバチ、カリバチ、アリなど）に限定して述べる。

膜翅目はいわゆる 1 倍体・2 倍体生物 (haplo-diploid) であって、雌は 2 倍体 (2 組の染色体を持つ) であるが、雄は未受精卵から発生するので雌親の染色体を 1 組だけ持つ 1 倍体である。雌親には雄と雌の産み分けが可能であるらしい。たとえばミツバチの場合、女王は結婚飛行の交尾で得た多数の精子を貯精嚢に貯え、これを生涯にわたって使う。貯精嚢は輸卵管のそばにあって、卵が卵巢から下ってきたとき、貯精嚢頸部の括約筋がゆるむと精子が袋から出て卵に達し、受精が起こって卵は雌となるが、括約筋が収縮すると卵は受精されず雄となる。一方、ワーカーの作る巣室は雄用と雌用で大きさが違うが（雄用の方が大きい）、女王は産卵に先立ち前肢で巣室の大きさをさぐり、大きい巣室には未受精卵を、小さい巣室には受精卵を産みつける。以上のことから、雄と雌の産み分けの行われていることが察せられる。

さて通常の 2 倍体生物では、前述のように母と娘の間の血縁度は同胞間と同じで $1/2$ であるが、膜翅目では上記の繁殖様式によって血縁度に差が生じる。母と娘の間の血縁度は $1/2$ であるが、完全同胞の姉妹間の血縁度は $3/4$ になる。これは次のように説明される。ある遺伝子が母親経由で姉妹間に共

有される確率は通常の 2 倍体生物と同じで $1/2 \times 1/2 = 1/4$ であるが、父親経由の場合、父親は 1 倍体であるから、その遺伝子は全部そっくり娘に伝えられる。したがって父親経由で姉妹が同一遺伝子を共有する確率は $1/2 \times 1 = 1/2$ である。母親経由の場合とあわせて、姉妹間で同祖遺伝子を共有する確率、すなわち同父母の姉妹間の血縁度は $1/4 + 1/2 = 3/4$ となる。つまり 1 倍体・2 倍体生物では、姉妹間の血縁度は母娘間より 50% 高いことになる。ここから結論されることは、ワーカーとして生まれた膜翅目の雌にとっては、生まれた巣を去って新たに巣を作り子を産んで育てるよりも（ワーカーが不妊であるのは女王の分泌する女王物質の作用で卵巢の発育が抑制されるためのようで、実際、女王を除くと妊性を回復することが多い）、巣に止まって妹を育てた方が遺伝子にとって有利だということである。このようにしてワーカーというカストの進化が説明される。

この説明で一つ問題になる点は、ミツバチなどの場合、女王が結婚飛行で交尾する雄は通常複数（7～10 匹）なので、巣には完全同胞ばかりでなく相当数の半同胞が含まれているという事実である。半同胞の姉妹間の血縁度は $1/4$ であるから、完全同胞だけを選んで世話をする場合（この可能性は薄い）を除いて、姉妹間の血縁度の平均は $(3/4 + 1/4) / 2 = 1/2$ となり、母娘間と同じになってしまう。この疑問はまだ完全には解けていないが、Trivers & Hare (1976) は結婚飛行の交尾が 1 回限りのアリを対象にして、上記の推論を間接的に立証しようとした。やや混み入った議論を必要とするが、あとで述べるゲーム理論的な考え方が使われていて興味深いので簡単に説明する。

ワーカーにとって妹との血縁度は上述のように $3/4$ であるが、弟との血縁度は $1/4$ である。何故なら、弟は未受精卵から発生して父親を持たないので、ワーカーが弟と共有する同祖遺伝子は母親経由のものに限られるからである（なお、弟からみた姉の血縁度は $1/4$ ではなく $1/2$ である。このような血縁度の非対称性からも興味ある推論が生まれる）。そこでもしワーカーが妹と弟を無差別に育てたとすると、血縁度の平均は $1/2$ となって自分の子を育てたのと違いがなくなってしまう。したがってワーカーが幼虫の性を見

分けられるときには、ワーカーは雄よりも雌を余計に育てた方が自分の（子育ての）遺伝子にとって有利だということになる。一方、女王の立場で考えたらどうであろうか。女王は息子にも娘にも遺伝子の半分以上を伝えるから、息子との血縁度も娘との血縁度も $1/2$ であり、雄と雌とを差別する理由がない。このように巣の幼虫の望ましい性比に関してはワーカーと女王の利益が一致しないのである。女王は前述の仕組みで雄と雌を産み分けることができるが、産まれた子の世話はすべてワーカーにまかされているので、ワーカーはその世話（とくに給餌）を通じて幼虫の生殺与奪の権を握っている。つまり女王の産み分けの性比が極端（たとえば雄ばかり）でない限り、ワーカーと女王はそれぞれある程度、巣の性比をコントロールする力を持っているといえる。そこで巣の性比をめぐる女王とワーカーとの間に争い（もちろん進化史上での）が存在したと考えると、この争いの結果はどちらに有利であったろうか。もし女王が勝ったのなら、巣の性比はほぼ $1:1$ となるはずであり、ワーカーが勝ったのなら、ほぼ雄 1 : 雌 3 の割合となるはずである。どちらであるかは実際に巣の性比を調べてみれば分かる。

しかしその前に、ワーカーにとって有利な性比が何故 $1:3$ で $0:1$ でないのかを説明する必要があるだろう。ワーカーにとって雄より雌を育てた方が有利なら、雄を育てることを止めて（雄への給餌を止めればよい）雌ばかり育てた方がよいと考えられるからである。これは次のように考えれば納得がゆく。膜翅目に限らず一般に有性繁殖をする生物にとって、繁殖集団の雄・雌の性比は繁殖成功度と密接に関係している。繁殖集団の性比が雌に偏ると、雄はより多くの交尾機会を持つことになり、雄の方が雌よりも 1 匹あたり多くの子をつくる。そこで親が生まれてくる子の性比をコントロールできる場合、親は娘より息子を多く産んだ方がより多くの孫を残せることになる。その結果（この場合、親のダーウィン適応度を孫の数で表わすことができるので）、性比を雄に偏らせるような淘汰圧が働くだろう。逆に性比が雄に偏ると、雌は一般に繁殖期に産む子の数が限られているので交尾機会が増しても得をしないが、雄の間で交尾をめぐる競争が激しくなるため、雄のつくる子の

数は雌にくらべて少なくなる。そこで今度は、性比を雌に偏らせるような淘汰圧が働くことになる。このように〈頻度依存性淘汰〉（この場合の頻度は性比である）が働く結果、性比は雄と雌の繁殖成功度が等しい（親にとっては、息子のつくる孫の数と、娘のつくる孫の数とが等しい）ところに落ち着くことになる。この性比は通常 1:1 であるが、多くの生物で雄・雌の性比がほぼ 1:1 であるという事実は、上のような頻度依存性淘汰の働きを推察させる。ただし人間を含む多くの高等生物の親は産まれてくる子の性比を制御できず、性比がほぼ 1:1 であるのは、配偶子形成に際しての性染色体のランダムな分離と、受精に際してのランダムな結合による。ここで考えられるのは、このランダム機構そのものが、1:1 の性比を実現する手段として進化したのではないかということであるが、この推測を実証することは容易でない。そこで生物界に時に存在する異常性比（1:1 の性比からのずれ）が上記の理論の検証材料として興味をひくこととなったのである。

さて膜翅目のワーカーにとって妹の血縁度は弟の 3 倍であった。よって弟の繁殖成功度が妹の 3 倍あれば、ワーカーにとって両者経由の遺伝的利益が等しくなって安定することになる。そのためには雄 1:雌 3 の性比であればよいことが上記の議論から理解できるだろう。これがワーカーにとって有利な性比が 1:3 である理由である。同様に、女王にとっての有利な性比は 1:1 であることが理解される。

実は、もう一つ考慮しなければならないことがある。それは有利な性比というとき、単に雄と雌の数の比ではなく、親にとっての投資量の比を考えねばならないということである。親は一定の資源（子をつくり育てるために親が利用できるさまざまなエネルギーや時間の総称）をそれぞれの子をつくり育てるのに配分して投資すると考えると、有利な性比とは息子と娘への有利な投資配分に他ならない。いま雄 1 匹をつくり育てるのに必要な資源の量が雌 1 匹の場合の n 倍であったとしよう。ということは親にとって、息子経由の孫 1 匹は娘経由の孫 1 匹の n 倍のコストがかかることであるから、 $n > 1$ のときは娘を多く産んだ方が有利である。ところが繁殖集団の性比が雌に偏

るにつれて上述のように雌の繁殖成功度が下ってゆくから、娘経由の孫1匹のコストが上昇する。バランスする点は、孫1匹のコストが娘経由と息子経由で等しくなる点、すなわち雄対雌の性比が数の上で $1:n$ になるときである。このとき親にとって、息子全体への投資量と娘全体への投資量の比は $1:1$ になっている。 $n \leq 1$ の場合も全く同じである。

以上の議論は、親にとって孫1匹の価値が息子経由でも娘経由でも等しいことを前提にしているが、膜翅目のワーカーの場合、たびたび述べたように、妹の子の価値は弟の子の価値の3倍であるから、投資量の比も弟1に対して妹3でなければならないことが容易に理解されよう。もちろん女王にとって好ましい投資比は $1:1$ である。

さて実際はどうであろうか。親の投資量はほぼ子供の大きさ（水分を除いた重量）に反映すると考えると、投資量の比は雄と雌の biomass（生物量、一定の地域に生存する生物の総乾燥重量）の比として捉えることができる。Trivers & Hare は 21 種のアリについて、この biomass の比を調べたところ、21 種のほとんどが $1:1$ より $1:3$ （雄対雌）に近いことを確かめた。つまり前述の、巣の性比をめぐるワーカーと女王の進化の争いではワーカーが勝利を収めたことになるのである。しかしながら、彼らのデータにはなお相当のバラツキがあり、また結果の解釈についての異論も存在する (Alexander & Sherman 1977)。またアシナガバチでは $1:1$ に近い投資比の報告もあるなど (Metcalf 1980)、問題ははまだ未解決の部分が多く含んでいる。

一方、〈Hamilton の規則〉 ($rb - c > 0$) を直接検証しようとする試みもいくつかあり、とくに鳥のヘルパー（ヤブカケスの若鳥などが、成熟後もしばらく繁殖を始めず、親のなわばりに止まって親の子育てを助ける現象）についての研究が活発であるが、 r と b と c の実測上の困難もあって明確な結論を得るに至っていない（総説として Emlen 1984 など）。このように実証的にはなお多くの問題が残されてはいるが、Hamilton の血縁淘汰の理論が Darwin の示唆を具体的に〈利己的な遺伝子〉の理論として展開し、生物の

協力行動の研究に新生面をひらいた功績は甚だ大きいと言わねばならない。

さらに、もう一つの理論的に重要な問題は、上述の性比(投資比)の議論のなかにゲーム理論の発想が含まれていることである。そもそも性比の進化の理論は前記の Fisher (1930) に由来するが、ゲーム理論の登場はこれよりあとである (von Neumann によるミニマックス定理の証明は 1928 年にさかのぼるが、ゲームの理論が体系的理論として提示されたのは 1944 年出版の Morgenthern との共著 "*Theory of games and economic behavior*" においてである。なお後述の効用の公理系は 1947 年の第 2 版から Appendix として付け加えられた)。しかし上掲書において Fisher はゲーム理論の基本を先取りしていたと言ってよい。前述のように、ある親による性比の選択の結果は他の親の選択によって変化する。たとえば雄の数が雌にくらべて少ないとき、雄を産んだ方が親の利益は大きい、すべての親が雄を産めば集団の性比は雄に偏り、その結果、雄を産むことの利益は減少する。そこで、ある親にとっての最適な性比の選択は他の親の選択を考慮して決められなければならないことになる。このように、自分の利益が自分の選択と同時に他者の選択によっても左右されるというのがゲームの理論の想定する場面であり、ゲームの理論は一般にこのような状況で、どのような選択の組み合わせが均衡するかを問題にする。Fisher はゲームの理論を知らなかったが、彼の問題にした均衡性比(投資比)は、その後ゲームの理論が一般化した非協力ゲームにおける Nash 平衡解に等しいのである。

元来ゲームの理論は上掲の著書の標題が示すように、経済行動の分析を目的としてつくられたために、人間の知的で意識的で目的志向的な行為を暗示するような用語が多い。ゲームという言葉自体がそうであるが、選択主体をプレイヤーと呼び、選択の行動経路を戦略 (strategy) と言い、選択の結果を pay off (支払。利得ともいう) と呼ぶなど、高度な知的活動を連想させるニュアンスがある。しかしゲームの理論自体はさまざまな状況での平衡解に関する数学的な定理の集りであって、定義の条件さえみたされれば生物現象にも無生物現象にも適用できる。決定者に知性(計算能力)や意識の存在する

ことは、この理論を適用するための必要条件ではない。たとえば平衡解は計算によって求めることができるが、決定者に計算能力がなくても、適当なフィードバック機構（決定者の選択に伴う利得の変化が決定者の将来の選択に影響するような機構）が存在すれば、平衡解が自動的に達成される場合がある。実際、生物の生存競争がダーウィン適応度を利得とし生物の形態・行動等を戦略とするゲームであると考え、自然淘汰はまさにこのようなフィードバック機構にあたり、長い進化の過程で平衡解を自動的に実現することになる。どのような平衡解が実現するかはゲームのルールと条件（ゲームの利得行列、コミュニケーションの有無、提携の可能性など）によるのであって、生物個体の意図や意識（そのような心的機構を想定できる場合でも）とは直接かかわりがない。これは人間の場合も同じである。人間はもちろん複雑な知的計算に基づいて行為を決定することができるが、結果が計算通りになるとは限らない。とくに多数の人間の相互作用が結果を左右するときには計算通りにならない方が普通である。結果を決めるのは結局、場面のゲーム的狀況（それは時に分析不可能なほど複雑なこともあるが）であるとすれば、知的計算能力の有無高低にかかわらず、人間と動物、時には植物をも一貫した視点で扱う道が開けることになる。元来、ゲームの理論はこのような意味での広い相互作用の理論であったのであり、そのことはすでに Fisher の前掲書の中に暗示されていたのであるが、実際に進化のゲーム理論として展開されるにはなお 40 年余りを必要とした。それは 1970 年代のはじめ、イギリスの進化生物学者 Maynard Smith らによって始められ、今日なお活潑に進められている（最近までの成果のまとめとして Maynard Smith 1982 がある。文献も同書に詳しい）。進化のゲーム理論を全般的に論ずることは別の機会にゆずり、ここでは専ら Darwin の提起した問題にかかわる部分だけに集中して論じたい。

4 互惠的利他行動と社会的ディレンマ

人間を含む多くの生物は、血縁でない間でも協力行動をとることがしばしば

ばある。この一見、利他的な行動の進化を利己性から説明するために Darwin の提案したのが互惠主義（相互扶助）の原則であり、それを詳細に展開したのが Trivers の理論であることをすでに述べた。Darwin や Trivers の理論は常識的に分かりやすいが、その最大の難点は、たびたび述べたように裏切り・ただ乗り防止の理論を欠くことである。たとえば雄と雌の番いによる子育ての場合、番いが協力すれば子の生存率は一般に上昇するから、両親とも協力によって適応度の上で利益を得るが、相手に子育てを押しつけて自分は新しい配偶者を見つけ、これとの間に新たに子をつくることができれば、適応度の上で一層の利益が得られる。実際、生物界には、両性が協力して子育てにあたる場合と、何れか一方の性が専ら子育てにあたる場合とがあるが、互惠的利他行動論ではこうした違いをうまく説明できない。

便乗の利益が存在するとき、相互に相手への便乗を計れば相互非協力となり、相互協力よりも悪い結果になる。自己の利益を求める行動が、かえって自己の利益を損うことになるのであるから、これは一種のディレンマである。また複数個体の相互作用の結果として生じるディレンマという意味で社会的ディレンマといってよいであろう。ここで社会とは、小は2個体の関係から大は地球的規模にまで及ぶ。Hobbes の指摘した「各人の各人に対する戦争」のディレンマはその古典的な例であるが、最近になって、現代社会の抱える地球的規模の問題の多くにこのような社会的ディレンマが含まれるという指摘が各方面からなされるようになった。たとえば過剰人口、資源涸渇、環境汚染の問題は、何れも多数の主体が協力（人口抑制、資源保護、環境保全）の利益を知りつつも、便乗することの一層の利益を求めて、あるいは便乗されることの不利益を怖れて行動する結果の社会的ディレンマであるという指摘、また大国間の軍備競争は少数主体間の同様な社会的ディレンマであるという指摘などが数多くなされている（たとえば Hardin 1968, Platt 1973, Dawes 1980 など）。実はこうした個別の指摘に止まず、もっと広く、厚生経済学から公共経済学へ、さらに公共選択の理論へと展開する最近の経済学と政治学の動き全体が深く社会的ディレンマの問題とかがかかわっている

が、これは冒頭に述べたように、Hobbes と Adam Smith に始まる近代政治学と経済学の潮流の必然的な帰結であるとみることができるのではないだろうか。

以上のように社会的ディレンマの問題は生物学、心理学、経済学、政治学にまたがる大問題であるが、以下、生物学と心理学と政治学とが接する地点から、社会的ディレンマの一つの典型である〈囚人のディレンマ〉に接近する試みに限定して考察したい。

社会的ディレンマの一番簡単な場合として2個体間の相互作用を考え、さらにディレンマを鋭くするために、相手が協力する場合はもちろん、協力しない場合も、自分は協力しない方が利益が大きいとすると、これはゲームの理論で有名な〈囚人のディレンマ〉(Prisoner's Dilemma. 以下 PD と略称)となる(この命名の由来等については Luce & Raiffa 1957)。

	<i>C</i>	<i>D</i>
<i>C</i>	<i>R, R</i>	<i>S, T</i>
<i>D</i>	<i>T, S</i>	<i>P, P</i>

これを行列ゲームの形で表わすと図のようになる。*C*は〈協力〉(Cooperate)の選択を、*D*は〈非協力〉あるいは〈離反〉(Defect)の選択を表わし、また各行は一方の個体(行プレイヤーという)の選択を、各列は他方の個体(列プレイヤーという)の選択を表わす。行列の各セルの二つの記号のうち先の記号は行プレイヤーの利得を、あとの記号は列プレイヤーの利得を表わす。たとえば行プレイヤーが *D* を、列プレイヤーが *C* を選択すると、行プレイヤーは *T* の利得を、列プレイヤーは *S* の利得を得ることになる。*R, S, T, P* はそれぞれ利得の数値を表わす記号であって単に *a, b, c, d* と書いてもよいところであるが、Rapoport の与えた意味づけが直観的に分かりやすいので、ここでもそれを用いる (Rapoport & Chammah 1965)。それによると、*R* は相互協力への報酬 (Reward) を、*T* は裏切り、あるいは便乗の誘惑 (Temptation) を、*S* は相手に裏切られるお人好し (Sucker) あるいは聖人 (Saint) の利得を、*P* は相互非協力の罰 (Punishment) を表わす。そし

て囚人のディレンマ PD の定義は (1) $T > R > P > S$ (2) $R > (T+S)/2$ の 2 条件である。

まず $T > R$, $P > S$ によって行プレイヤーも列プレイヤーも、相手の選択の如何にかかわらず、 D を選択した方が利得が大きい (ゲーム理論の用語でいえば D が C に〈優越する〉)。ところが双方が D を選ぶと、 $R > P$ であるから、結果は相互協力の場合に劣るというディレンマが生じる。上記の三つの不等式をあわせると条件 (1) となる。経済学の用語でいえば、 DD (双方 D という選択。以下同様) は均衡解であるが、パレート最適ではないというディレンマである。最初に述べたように、一般均衡理論は、ある一般的な仮定の下に、市場の価格機構を通じてパレート最適な均衡解が実現することを証明したが、同時にその仮定がみたされない場合のあることを明らかにした (〈市場の失敗〉〈外部性〉〈公共財〉など)。上の PD は また別の側面から、パレート最適という経済的効率性が実現されない場合のあることを示している。

さて PD を繰返して行ったらとき、プレイヤーが暗黙の了解の下に CD 選択と DC 選択とを交替させた (交互に裏切り合った) とすると、両プレイヤーの平均利得は $(T+S)/2$ となるが、この平均利得が相互協力の利得 R を超えないようにするのが条件 (2) である。この条件は 1 回限りの PD では必ずしも必要でないが、後に取り上げる繰返しのある場合のために設けておく。

ところでゲームの利得とは実際場面では何を表わす数値であろうか。進化ゲームでは、それはすでに述べたように、ダーウィン適応度 (の変化) であることがはっきりしている。実際、Maynard Smith (1982) はダーウィン適応度 (の変化) を利得とする行列によってさまざまな進化ゲームを表現し、またダーウィン適応度によって進化における一種の平衡戦略を定義している (Nash 平衡解をやや強めた定義であるが、突然変異によって侵略されないような集団の支配的な戦略という意味で Maynard Smith は *Evolutionarily Stable Strategy* (進化的に安定な戦略)、略して *ESS* と呼んでいる)。ダーウィン適応度はたびたび述べたように生物が次世代に残す子孫の数であるか

ら、それが数値で表わされるのは当然である。

一方、人間社会の社会的ディレンマにおいて争われる利得は直接的にはダーウィン適応度ではない。人間も進化の産物である以上、究極的にはダーウィン適応度の支配下にあると言ってよいが、進化のゲームでは地質学的な長い時間をかけて勝敗が争われるのに対して、人間社会の社会的ディレンマでは地質学的には瞬間に等しい短時間の間の相互作用が問題になる。そこでの利得は経済学とゲーム理論の枠組で言えば、一般に〈効用〉(utility)と呼ばれるものになろう。効用とは、ゲームの結果の好ましさをプレイヤーごとに評価した数値である。プレイヤーが個人であれば個人的評価であるが、国家のような個人の集団であれば、集団成員の個人的評価を何らかの仕方で集計した社会的評価となる。ここにはダーウィン適応度の場合と違って、少なくとも二つの困難がある。その第一は〈効用の可測性〉の問題である。周知のように、この問題は近代経済学の歴史とともに古い。一方において序数効用のみを仮定する〈無差別曲線〉に基づく理論があり、他方において von Neumann (前述 1947 の第 2 版 Appendix) の公理系に基づく (あるいはそれを若干修正した) 基数効用の理論がある。またこの問題には心理学も深くかかわっていて実験的研究も数多く行われているが、いまだに明確な結論を得るに至っていない (なお PD の条件 (1) には序数効用で十分であるが、あとで触れるような理論的計算には基数効用 (間隔尺度) の仮定が必要となる)。第二の困難は、プレイヤーが集団であるとき、集団成員の個人的評価を合成して妥当な社会的評価をつくることからなる問題である。個人的評価と社会的選択との間には、冒頭に触れた Arrow (1951) 以来の困難なパラドックスが存在するし、また効用の個人間比較の可能性の問題もあって、何れも数多くの研究にもかかわらず、まだ十分解明されるに至っていない。

このように困難な問題は残されているが、個人的評価であれ社会的評価であれ、何らかの評価の繰返しがいわば社会的(文化的)淘汰として働いて、こうした淘汰作用の積み重ねが文化をつくり、また既成の文化を変えてゆくという、一種の文化的進化を考えることが許されるならば、ダーウィン適応度

と効用の間に、ある種の関係をつけることができるのではないだろうか。それは両者の何れも、ゲーム（生物学的進化あるいは文化的進化の）の戦略の、世代を通じての再生産に関係しているという見方である。生物学的進化における戦略頻度の世代間変化とダーウィン適応度の関係は、遺伝法則と集団遺伝学のモデルを使って求めることができる。それは時に複雑な関係になることもあるが、多くの場合、両者の間には単調増大の関係がある。つまり平均してダーウィン適応度の高い戦略ほど繁殖集団内でその頻度を増してゆく。同様にして、文化的進化における戦略頻度の世代間変化と効用の関係を求めるには、遺伝法則に相当する文化伝承の法則と、集団遺伝学に相当する数理的モデルが必要であるが、そのような法則やモデルはまだ発見されていない（若干の興味ある試みは存在する。たとえば Cavalli-Sforza & Feldman 1981）。したがって正確な議論をすることはできないが、大雑把に言えば文化的進化の場合も、平均して効用の高い戦略が文化集団内で頻度を増してゆくと考えてよいのではないだろうか。何故なら、効用という評価の機構自体も文化の淘汰作用を受けるはずであって、戦略の再生産能力を正しく反映しないような評価機構は次第に排除されるだろうと思われるからである。このような文化的淘汰のつくり出す効用は、戦略の再生産頻度の反映として基数尺度であり、また戦略使用者個人を超えた評価として個人間比較の可能な尺度であるとみなしてよいだろう。ただしそれをどう測定するかという問題は将来に残されている（基本的には世代間の再生産能力を直接測るという方向もある）。以上の考察に基づき以下の議論では、ダーウィン適応度と効用とを区別せずにとともに利得として、また戦略の再生産能力を反映する指標として考えてゆくこととする。

さて PD に戻って考えよう。このゲームでは DD 選択が唯一の平衡解として進化的にも文化的にも安定であることは明らかであるが、それではこのディレンマから脱して CC 選択を解として成立させる方法としては、Hobbes 流の外部からの強制しかないのであろうか（プレイヤーが相談の結果、CC 選択に合意した場合も〈裏切り〉〈ただ乗り〉を防ぐには外部からの強制が必

要である)。外部からの強制とはつまり、*DD* 選択に罰という形の負の利得を加え、*CC* 選択に賞という形の正の利得を加えて、*PD* を非 *PD* に変換することであるから、これでは *PD* 自体を解決したことにはならない（人間社会の多くの社会的ディレンマが外部からの強制によって解決されていることは事実であろうが、これは別問題である）。

一つの解決法は、社会的ディレンマの多くが1回限りで終るゲームではなくて、同一のプレイヤーの間で、ほぼ同一の条件で何回も繰返されることに着目して、*PD* を繰返しのあるゲームとすることである。繰返しのあるゲームでは、プレイヤーは選択を通じて自分の意志を相手に伝え、その後の相手の選択に影響を与えることができる。つまり1回限りの場合と違って、選択を通じてのコミュニケーションが可能になるので、遙かに複雑な戦略を構想できるようになる（なお繰返しを入れた場合も〈非協力〉ゲームとして、選択反応以外の手段をもってするコミュニケーションは禁じられているとする）。

ただし繰返しがあっても、その回数があらかじめ決まっている場合は、1回限りの *PD* と本質的に同じであることが次のようにして示される。最終回の選択はもはや相手への影響を通じて相手からの応答を受けることがないから、その回で最も有利な選択である *D* が選ばれることになる。ところがこのようにして最終回の選択が決まると、その一つ前の回が事実上の最終回となるから、その選択も *D* に決まる。以下同様にして一つずつ前の回が事実上の最終回となり、遂に1回目までさかのぼることになって、結局、1回目から最終回まですべて *D* という反応系列が唯一の平衡解となるのである。

これを避けるには、あらかじめ確定した最終回の存在しない——つまり毎回、正の確率で持続する——ゲームを考えればよい。その最も簡単な場合として、毎回、ある一定の正の確率 ($w > 0$) で繰返す *PD* を考え、これを〈繰返しのある *PD*〉 (Iterated Prisoner's Dilemma) と呼び、IPD の略称で表わすことにする。また w を〈繰返し確率〉と呼ぶ。IPD の長さ l の分布は、 n を任意の正の整数として

$$\begin{aligned} Pr\{l \leq n\} &= 1 - w + w(1 - w) + w^2(1 - w) + \cdots + w^{n-1}(1 - w) \\ &= 1 - w^n \end{aligned} \quad (1)$$

であり、この分布関数を使うと、 l の平均 \bar{l} は

$$\bar{l} = \frac{1}{1 - w} \quad (2)$$

となり、また l のメディアン \tilde{l} は $Pr\{l \leq \tilde{l}\} = 1/2$ から

$$w^{\tilde{l}} = 0.5 \quad (3)$$

$$\tilde{l} = -\frac{\log 2}{\log w} \quad (4)$$

となる。たとえば $\tilde{l} = 200$ なる w は $w = (0.5)^{0.005} \doteq 0.99654$ であり、また $w = 0.5$ のとき $\tilde{l} = 2$, $\bar{l} = 1$ である。

さてこのように定義された IPD について *All D* (D の連続選択) 以外に妥当な解が存在するだろうか。また存在するとすればそれは何であろうか。政治学者の Axelrod はこの問題に実証的に答えるために、一つの簡単な IPD ($T=5$, $R=3$, $P=1$, $S=0$ としたもの) について広く各方面から募集したプログラム形式の戦略をコンピュータ上で争わせ (コンピュータ・トーナメント), その結果を分析した (Axelrod 1980a, 1980b, 1984)。またこの分析結果を裏づけ、かつ一般化する理論を展開した (Axelrod 1981)。さらに前述の血縁淘汰理論の Hamilton との共著論文 (Axelrod & Hamilton 1981) ならびに著書 (Axelrod 1984) では、この理論の持つ政治学的意味と生物学的意味とが考察されている。以上の IPD に関する実証的・理論的研究は、方法論的にも内容的にも示唆するところが多いので、以下にその概略を論評する。

まずコンピュータ・トーナメントについてであるが、従来、PD の実証的研究といえば、実験室で被験者 (多くは大学生) をペアにして利得行列を示し (ペアの一方がサクラ, すなわち実験者があらかじめ設定した戦略の実行者であることもある), 相互に独立に繰返し選択させるという、〈実験ゲーム〉 (experimental game, gaming) が圧倒的多数であった。Rapoport & Chammah (1965) の組織的研究をはじめ、PD 実験ゲームの報告は現在ま

で膨大な量に達するが、研究結果の与える情報量は意外に少ない（最近までの研究の論評として Pruitt & Kimmel 1977 がある。文献も詳しい）。実験結果は通常、被験者の *C* 反応比率、相互協力反応 *CC* の比率、相互非協力反応 *DD* への閉じ込め率、サクラによる戦略操作の効果、各種反応比率の系列依存性や条件（被験者の性別、パーソナリティなど）別変化など、詳細にわたって分析されているが、一般に *C* 反応比率などが試行の進行とともに多少とも規則的な変化を示すこと、サクラ戦略のうちでは後述の *TFT* が *C* 反応を誘いやすいことなどを除いて、意味深い結果はあまり得られていない。その理由としては実験場面と実生活場面との違いなどいろいろ考えられるが、最大の理由は次のような事情にあるのではないかと思われる。本来、ゲームの理論でいうゲームとは戦略（行動プログラム）の争いである。しかし実験ゲームの被験者は試行ごとの対応に追われ、試行を通じての一貫した戦略を争う余裕もなければ、また戦略を構想する十分な動機づけも持たないのが普通であろう。ただし意識的な戦略構想がない場合でも、前述のフィードバック機構が働けば、結果的に戦略が成立する可能性はあるが、通常の実験ゲームの利得構造（被験者の獲得した得点は実験終了後、金額に換算して支払われることが多いが、その金額は比較的少額である）や繰返し数（実験時間の制約もあってそれほど多くない）を考えると、十分有効なフィードバックを期待することは無理であろう。

一方、ゲーム場面における行動プログラムはコンピュータのプログラムとして明示することができる。日常の行動プログラムはもっと曖昧であってコンピュータ・プログラムのように明示的でないという批判に対しては、プログラムに適当にランダム要素を加えて曖昧化することはむしろ容易であると答えることができよう。被験者に利得行列を示して、実際に相手と争う前に、あらかじめいろいろな場合（さまざまな相手とその戦略）を想定した上で最良と思われる行動プログラムを書かせれば、ゲームの理論でいう戦略の実態が従来の実験ゲームよりも、よく捉えられるのではないだろうか。その意味で Axelrod がコンピュータ・プログラムの形で *IPD* の行動プログラ

ムを募集し分析したことは、ゲームの実証的研究に有望なアプローチを開いたといえることができる。

プログラムの募集は2回にわたって行われた。1回目の募集は心理学、社会学、経済学、政治学、数学の分野でゲームの理論ないし実験的研究の業績のある者を対象にして行われたが、全部で14のプログラムの提出があった。なおゲームの行列は前記の通りであるが、応募者にはあらかじめ、提出された戦略は総当たり制で(自分自身との試合も含む)、各ペアにつき1試合200試行で5試合ずつ戦わせることを知らせてあった。通常の実験ゲームにくらべて試行数が多いこと、また相手との組み合わせの多様なことが注目されるが、このことはおそらく、応募者に戦略的発想をとらせる上で効果があったと思われる。なお試行数が確定していたことは前述のように理論的には問題があり、実際、提出された戦略においても最後の数試行で *D* が連続して使われる傾向があったので、結果の分析は最終の数試行を除いて行われた。また2回目の募集は後述のように、試行数不確定として行われた。

提出された14の戦略にランダム戦略を加えた計15のプログラムのコンピュータ・トーナメントの結果(試行総数は $15 \times 15 \times 200 \times 5 = 225,000$)、きわめて明瞭な傾向が浮かび上がった。それは第一に、全試合を通じて最高点を獲得したのは、15のプログラムのうちで最も簡単な *TFT* 戦略であったことである。この戦略はゲームの実験的研究で著明な Anatol Rapoport の提出したもので、*TFT* とは *TIT FOR TAT* (しっぺい返し) の略であるが、その内容は頗る簡単であって、①第1試行は必ず *C* を選択し、②第2試行以降は一つ前の試行での相手の選択と同じ選択をする、というものである。たとえば相手の選択の系列が *DCCDCD*...であれば、*TFT* の選択は *CDCCDCD*...となる。第二の明瞭な傾向は、*TFT* を含め一般に〈善良な〉(nice) 戦略(相手より先に *D* を選択することは決してしない戦略の総称)が非善良戦略より好成績をあげたことである。

さて以上のような第1回目の試合の結果を知らせた上で第2回目の募集が行われたが、今度は6ヶ国から合計62の応募があった。応募者の専門は1

回目の5分野の他に物理学、生物学（前記の Maynard Smith）、コンピュータ科学に及び、また多数のコンピュータ hobbyist（最年少は11歳の少年）を含んでいた。1試合の試行数は今度は不確定で、 $w=0.99654$ のIPDであった（このとき前述の(4)式から試行数のメディアンは $\bar{l}=200$ となる。実際の試行数の選択は、 $\bar{l}=200$ とした(1)式の分布からのサンプリングによった）。コンピュータ上の試合は、62のプログラムにランダム戦略を加えた63の戦略の間で総当たり制で（自分自身との試合も含む）行われたが、約300万試行からなるデータは再び明瞭な傾向を示した。重ねて Rapoport の提出した *TFT* が最高点を獲得したのである。他のプログラムの多くは、1回目の試合における *TFT* の好成績を強く意識し、さまざまな工夫によって *TFT* を改善し、あるいはこれを出し抜こうとするものであり、中には相手の反応系列をマルコフ過程で近似して自己の最適な条件つき確率を計算するなど相当に複雑な戦略もあったが、何れも最も簡単な *TFT* に及ばなかったのである。また1回目と同じく今回も〈善良〉戦略が上位を占めたが、善良戦略の間で優劣を決めるのは戦略の〈返報性〉(provocability)であることが示唆された。*TFT* は相手より先に *D* を選択することはないが（善良性）、相手が *D* を出せば直ちに（次の試行で）*D* を返すという、すばやい返報性がある。これを欠いた戦略は一般に *TFT* より劣る。たとえば Maynard Smith の提出した戦略は *TIT FOR TWO TATS* (*TFT* との違いは、相手が2回続けて *D* を出したとき初めて *D* を返すという点にある)であったが、総得点の順位は24位に止まった。その理由は、他に喰い逃げ戦略（相手の反応を窺って時々 *D* を出すが、2回続けて *D* を出すことはしないというような戦略）をとるプログラムがあって、これに徹底的に搾取されたためである。一方、*TFT* はこのような戦略に搾取されることがない。同時に *TFT* は、相手が *D* のあと *C* にもどせば、自分も直ちに *C* を返すという〈寛容性〉(forgiveness)を持っている。

Axelrod はさらに、63の戦略の生き残り競争をコンピュータ上で検討した。前述のように、ゲームの得点が、ダーウィン適応度と同様、基本的には

戦略の再生産能力を表わすと解釈すれば、そこから戦略の文化的進化を予測することができる。Axelrod は最も単純に考えて、トーナメントの得点そのものが、各戦略ペアの間で次世代に生き残る戦略の数を表わすと仮定して、世代の推移によって戦略の相対頻度がどう変わってゆくかをコンピュータで計算した。全部で 1000 世代の推移を通じて再び *TFT* の最優位が示されたが、ここに興味深いのは搾取的な戦略のたどった運命であった。上述の喰い逃げなど搾取的な戦略は当初、搾取の成功によって頻度を増してゆくが、やがて搾取対象が淘汰されるにつれて搾取の対象を失い、自らも急速に消滅していったのである。

以上のように、2 回のトーナメントと進化のシミュレーションの結果を通じて *TFT* の優位性はきわめて印象的であったが、Axelrod はその理由を *TFT* の〈善良性〉〈返報性〉〈寛容性〉および〈明瞭性〉(clarity. きわめて簡単でかつ一貫した戦略なので、相手がすぐ戦略の特徴を理解できること)に帰している。

ところで *TFT* の優秀性は、たまたま 2 回のトーナメントに参加した戦略の間だけでなく一般に言えることなのであろうか。それを知る一つの方法はこのようなトーナメントを数多く重ねることであるが、理論的に *TFT* の優秀性を示すことができればさらによい。それには *TFT* をあらゆる可能な戦略の中で位置づける必要があるが、あらゆる可能な戦略を考慮するというのは一般的には非常な困難を伴う。しかし幸いなことに IPD の場合、*TFT* の極度の単純さによって、それが可能になるのである。Axelrod (1981) は 8 つの定理の証明を通じて *TFT* の理論的な特徴を明らかにした。彼の得た最も重要な結論の一つに、いったん *TFT* が集団の支配的な戦略となると、前記の繰返し確率 w が十分大きい限り、*TFT* は安定な戦略であって、他の戦略の侵略を受けないという定理がある。一般に、このような安定性を持つ戦略を、Maynard Smith の *ESS* にならって Axelrod は *CSS* (*Collectively Stable Strategy*. 集団的に安定な戦略) と名づけた (厳密にいうと *ESS* と少し違い、*CSS* には Nash 平衡の条件がそのまま用いられているの

で、*ESS* より弱い条件で成立する)。*All D* (*D*の連続選択)は常に *CSS* であるが、もう一つの重要な定理の示すところによると、*TFT* が *cluster* (相互作用が主としてその内部に限られているような部分集団のこと)をつくって攻撃すれば *All D* の支配する集団を侵略することが可能であり、しかも *TFT* 自体は *cluster* の侵略に対して *All D* のような弱点を持たない。このような理論的結果と前記の血縁淘汰の理論とをあわせ考えると、次のような協力行動進化のシナリオをえがくことができる。

All D の安定支配 → 血縁小集団における *TFT* の発生 → *TFT cluster* による *All D* 集団の侵略 → *TFT* の安定支配 (相互協力)

Hamilton は前記の Axelrod との共著論文の中で、このようなシナリオによる生物進化を論じ、また *TFT* に相当すると思われるいくつかの生物学的事例をあげているが、このシナリオはまた人間社会の文化的進化にも適用できる可能性を持っている。

このように〈囚人のディレンマ〉をめぐる最近の研究は、①新たにコンピュータによるトーナメントとシミュレーションという方法を導入したこと、②ゲームの理論に新しい展望を開いたこと、③生物学者と社会科学者の協力を通じて生物学的進化と文化的進化とを共通の理論的枠組の中で論じる可能性を示したことなど、注目すべき展開をみせつつある。もちろん残された問題は数多いが、その中で特に次の三つを指摘したい。

1) ゲームの理論には一般に dynamic process の視点が不足しているが、IPD に関しても、たとえば *TFT* の支配拡大の過程を説明するような理論の発展がのぞまれる。

2) 社会的ディレンマとしては〈 n 人 ($n \geq 3$) 囚人のディレンマ〉の方がより一般的であると考えられるが、*TFT* の理論と実証を n 人ゲームに拡張できるかどうか。

3) ゲームの利得行列から戦略の成長と衰退を量的に予測するには、集団

遺伝学に匹敵する〈文化伝承のモデル〉が必要である。

これらは何れも大きな問題であるので、稿を改めて論じることとしたい。

引用文献

- Alexander, R. D. & Sherman, P. W. 1977. Local mate competition and parental investment in social insects. *Science* **96**, 494-500.
- Aoki, S. 1977. *Colophina clematis* (Homoptera: Pemphigidae), an aphid species with "soldiers". *Kontyû*, Tokyo, **45**, 276-82.
- 青木重幸 1984. 『兵隊を持ったアブラムシ』どうぶつ社.
- Arrow, K. J. 1951, (2nd edition) 1963. *Social choice and individual values*. Yale Univ. Press.
- Arrow, K. J. & Hahn, F. H. 1971. *General competitive analysis*. Holden-Day, Inc. & Oliver & Boyd. (福岡正夫・川又邦雄訳『一般均衡分析』1976, 岩波)
- Axelrod, R. 1980a. Effective choice in Prisoner's Dilemma. *Journal of Conflict Resolution* **24**, 3-25.
- Axelrod, C. 1980b. More effective choice in Prisoner's Dilemma. *Journal of Conflict Resolution* **24**, 379-403.
- Axelrod, R. 1981. The emergence of cooperation among egoists. *American Political Science Review* **75**, 306-18.
- Axelrod, R. 1984. *The evolution of cooperation*. Basic Books.
- Axelrod, R. & Hamilton, W. D. 1981. The evolution of cooperation. *Science* **211**, 1390-6.
- Cavalli-Sforza, L. L. & Feldman, M. W. 1981. *Cultural transmission and evolution: a quantitative approach*. Princeton Univ. Press.
- Darwin, C. 1859. *The origin of species by means of natural selection*. The Modern Library edition. (八杉竜一訳『種の起源』(上)(中)(下) 1963-1971, 岩波文庫)
- Darwin, C. 1871. *The descent of man and selection in relation to sex*. The Modern Library edition.
- Dawes, R. M. 1980. Social dilemmas. *Annual Review of Psychology* **31**, 169-93.
- Dawkins, R. 1976. *The selfish gene*. Oxford Univ. Press. (日高敏隆・岸 由二・羽田節子訳『生物=生存機械論』1980, 紀伊国屋書店)
- Dawkins, R. 1979. Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie* **51**, 184-200.

- Eldredge, N. & Gould, S. J. 1972. Punctuated equilibria—an alternative to phyletic gradualism. In T. J. M. Schopf (Ed.) *Models in Paleobiology* 82-115, Freeman, Cooper & Co.
- Emlen, S. T. 1984. Cooperative breeding in birds and mammals. In Krebs, J. R. & Davies, N. B. (Eds.) *Behavioural ecology: an evolutionary approach*, 2nd ed. 305-39, Blackwell Scientific Publications.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Dover edition, 1958.
- Gould, S. J. 1977. *Ever since Darwin*. Penguin Books. (浦本昌紀・寺田 鴻訳『ダーウィン以来』上, 下 1984, 早川書房)
- Grafen, A. 1982. How not to measure inclusive fitness. *Nature* 298, 425-6.
- Grafen, A. 1984. Natural selection, kin selection and group selection. In Krebs, J. R. & Davies, N. B. (Eds.) *Behavioural ecology: an evolutionary approach*. 2nd ed. 62-84, Blackwell Scientific Publications.
- Hamilton, W. D. 1964. The genetical evolution of social behaviour. I, II. *Journal of Theoretical Biology* 7, 1-16; 17-32.
- Hamilton, W. D. 1967. Extraordinary sex ratios. *Science* 156, 477-88.
- Hardin, G. R. 1968. The tragedy of the commons. *Science* 162, 1243-8.
- Hildenbrand, W. & Kirman, A. P. 1976. *Introduction to equilibrium analysis: variation on themes by Edgeworth and Walras*. North-Holland Publishing Co.
- Hobbes, T. 1651. *Leviathan*. Penguin Books edition, 1968. (水田 洋・田中 浩訳『リヴァイアサン』1974, 河出書房新社)
- Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* 217, 624-6.
- 木村資生 (編) 1984. 『分子進化学入門』培風館.
- Luce, D. & Raiffa, H. 1957. *Games and decisions*. Wiley.
- Macbeth, N. 1971. *Darwin retried: an appeal to reason*. Gambit. (長野 敬・中村美子訳『ダーウィン再考』1977, 草思社)
- Maynard Smith, J. 1982. *Evolution and the theory of games*. Cambridge Univ. Press.
- Metcalf, R. A. 1980. Sex ratios, parent-offspring conflict, and local competition for mates in the social wasps *Polistes metricus* and *Polistes variatus*. *American Naturalist* 116, 642-54.
- Packer, C. 1977. Reciprocal altruism in *Papio anubis*. *Nature* 265, 441-3.
- Platt, G. 1973. Social traps. *American Psychologist* 28, 641-51.

- Pruitt, D. G. & Kimmel, M. J. 1977. Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. *Annual Review of Psychology* 28, 363-92.
- Rapoport, A. & Chammah, A. M. 1965. *Prisoner's dilemma: a study in conflict and cooperation*. Univ. of Michigan Press.
- Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations*. The Modern Library edition, 1965. (大内兵衛・松川七郎訳『諸国民の富』(一)~(五) 1959-1966, 岩波文庫)
- Trivers, R. L. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46, 35-57.
- Trivers, R. L. & Hare, H. 1976. Haplodiploidy and the evolution of social insects. *Science* 191, 249-63.
- von Neumann, J. & Morgenstern, O. 1944. (2nd. ed. 1947) *Theory of games and economic behavior*. Princeton Univ. Press.
- Wright, S. 1969. *Evolution and the genetics of populations*. Vol. 2. *The theory of gene frequencies*. Univ. of Chicago Press.
- Wright, S. 1977. *Evolution and the genetics of populations*. Vol. 3. *Experimental results and evolutionary deductions*. Univ. of Chicago Press.
- Wright, S. 1978. *Evolution and the genetics of populations*. Vol. 4. *Variability within and among natural populations*. Univ. of Chicago Press.